Optimal Nonpoint Source Monitoring:

An Application to Redwood Creek

By

Richard E. Howitt

Department of Agricultural and Resource Economics

University of California, Davis

Davis, CA 95616

TECHNICAL COMPLETION REPORT

Project Number UCAL-WRC-W887

June 2000

University of California Water Resources Center

The research leading to this report was supported by the University of California Water Resources Center, as part of Water Resources Center Project UCAL-WRC-W-887.

ABSTRACT:

The research summarized in this report involves three interrelated analyses. First, we construct a theoretical nonpoint source (NPS) pollution control model and derive the optimal budget tradeoff between direct treatment of polluting sources versus data collection, which facilitates information acquisition and learning about the relative pollution loading among the sources. Second, we develop a sequential entropy filter to statistically update estimated NPS pollution loading parameters, as new data becomes available. We apply the entropy method to stream flow and ambient sediment loading data for Redwood Creek, which flows into and through Redwood National Park. Third, the theoretical and methodological results are incorporated into a sediment control model for Redwood Creek. We simulate the sediment control management program to provide policy analysis by comparing a uniform treatment policy where no data is collected against high and low-intensity data collection policies. The simulation results indicate that the high and low-intensity data collection policies can reduces uncertainty, increasing treatment effectiveness, so that sediment related damages are lower than damages under the uniform treatment policy.

Key words: Nonpoint Source Pollution, Optimal Monitoring, Water Quality Management, Maximum Entropy, Dynamic Modeling

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PROBLEM AND RESEARCH OBJECTIVES

Since the passage of the Clean Water Act in 1972, most pollution control measures have focused on the problem of reducing point source (PS) pollution. However, water resources in the United States are predominantly polluted by nonpoint sources (NPS) which are more difficult to measure (EPA 1995). For this reason, we choose to study NPS pollution control to better understand the economic underpinnings and the statistical nature of the problem.

Nonpoint source (NPS) pollution control is primarily an information, or conversely, an uncertainty problem, and as such, modeling NPS pollution control requires that the role of information and learning be explicitly specified and addressed in the model. In general, pollution control problems can be characterized by the degree of uncertainty or incomplete information about the location of polluting sources and the magnitude of their contribution. If there were complete certainty or perfect information on the location and magnitude of pollution for each source, then the problem would be a "pure" point source (PS) problem. If, however, the manager does not know the location of the sources (i.e., fully ignorant of the source locations), then the problem would be a "pure" NPS problem.[1]

These two extreme certainty scenarios mark the ends of a spectrum that defines all pollution problems by the degree of uncertainty about the location and pollution generation for each source rather than the vague classification of either PS or NPS. Note that the extreme NPS problem is not susceptible to policy controls because the location of the pollution sources needs to be known to implement treatment policies. That is, implementation of pollution control policies requires that the location of the sources is known with a high degree of confidence, or

---

[1]. If the manager cannot identify the sources, then it follows that she cannot know the magnitude of pollution generated by each source either. Even if she knows the location, due to the effects of a mix of random events, she may not be able to identify the contribution to the total pollution from each source.

information can easily be acquired about the identity of the sources so that treatment efforts can be undertaken.

Another complication to NPS control is the budgetary restrictions that often limit the extent of NPS pollution control. These financial limitations have implications for pollution control in general and particularly in the United States, where the federal government spends $3 billion annually to control NPS water pollution, yet reports that the public control of NPS pollution was limited by a lack of financial resources (GAO 1999). The United States government is perhaps the largest controller of NPS pollution and the largest contributor of NPS pollution, contributing nearly half of all the NPS pollution in the 11 western states (Arizona, California, Colorado, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, and Wyoming) (GAO 1999).

As an application of the theoretical and methodological aspects described above, the current sediment control program for Redwood Creek, which flows into and through Redwood National Park located in Orick, California, is considered. If perfect information were available on the location and sediment loading contribution of the pollution sources, park managers could allocate the entire sediment control budget to treatment effort. However, with incomplete information, the management of sediment loading requires an explicit or implicit allocation of sediment control resources between information acquisition and treatment. We focus on the tradeoff between information acquisition and treatment by explicitly modeling information acquisition and learning within a pollution control model. This pollution control model extends the model in Farzin and Kaplan (1999) to a dynamic problem.

The analysis of NPS pollution control is further complicated by the statistical nature of NPS pollution data. In any given period there are more polluting sources than observations (i.e., an undersized sample).[2] This statistical problem, sometimes referred to as an ill-posed data problem, makes traditional statistical approaches inappropriate and requires that a new approach be employed. In the empirical application to follow we adopt a sequential entropy filter (SEF) approach, which generates parameter estimates from ill-posed data and updates the estimates, as new data becomes available (Kaplan and Howitt 2000).

In the analysis two data collection regimes are constructed to facilitate the estimation of the unobservable sediment loading parameters. These estimates are then incorporated into the sediment control model. We simulate the model to derive optimal policy decisions regarding treatment and information acquisition strategies and provide a comparison of the current treatment policy with two alternative site-specific NPS pollution treatment policies. This comparison allows us to test the hypotheses that a flexible treatment strategy, which allows for site specific treatment effort, will lower sediment related damages if the manager can acquire information and shift resources toward dirtier pollution sources. The results drawn from this analysis can, in general, shed light on the role of information and budgetary constraints on the efficient management of pollution problems.

CASE STUDY

Redwood National Park was established in 1968 to preserve the coastal redwood (*Sequoia sempervirens*) forest ecosystem. The Tall Trees Grove, which includes several of the world's tallest trees, is located in the Park adjacent to Redwood Creek. Erosion from upstream

---

2. One could wait for a sufficiently long time series to be collected, thereby avoiding the ill-posed estimation problem. However, interim damage will occur and may result in irreversible losses, such as the destruction of critical habitat for endangered or protected resources. The information contained in the data may be incorporated as it is collected so that the manager can take action now and reduce the interim damage rather than wait until enough data is collected and face the risk of incurring irreversible losses. Farzin (1996) shows, theoretically, the importance of controlling interim damages.

logging activity and the associated road network increased sediment delivery to streams, resulting in destabilized stream channels, dramatic geomorphic changes, and the subsequent loss of aquatic and riparian habitat. Most importantly, the Tall Trees Grove was subject to increased flooding, bank erosion and an elevated water table (Spreiter et al. 1995).

To protect the grove, the United States Congress expanding the park in 1978 to include 36,000 additional acres upslope of the original Redwood Creek corridor. In 1981, Congress directed the United States National Park Service (DOI) to minimize erosion from past land uses, re-establishing native patterns of vegetation, and protecting aquatic and riparian resources within tributaries and along the main stem of Redwood Creek (Spreiter et al. 1995). The primary focus of erosion control would be to prevent or reduce erosion from logging roads within the Park (DOI 1981).[3]

To date nearly 100 road rehabilitation projects have been completed within the Park boundary. We collected data on treatment levels and costs, and data collection costs from unpublished project reports for the period spanning 1981 to 1988. This time period provides the most complete data set. Daily stream flow and ambient sediment load measures were obtained from United States Geological Survey (USGS) gaging stations located along Redwood Creek (See Figure 1). The Park arranged with the USGS to collect daily sediment load measures only at the Orick and Blue Lake gaging stations. The difference between the upstream and downstream sediment measures constitutes the total sediment entering Redwood Creek between Blue Lake and Orick. Daily stream flow, however, is measured at all seven gaging stations. In total, there are 1755 daily observations of stream flow and sediment loading spanning nine rain

[3]. Managers in the park identified the road network from past logging activity as the most probable sediment contributors to Redwood Creek. Logging roads are generally considered the principal contributor to sediment loading in logged forests (Mount 1995; GAO 1999; EPA 1999).

seasons. The data from the gaging stations delineates six catchment regions as polluting sources (see Table 1).

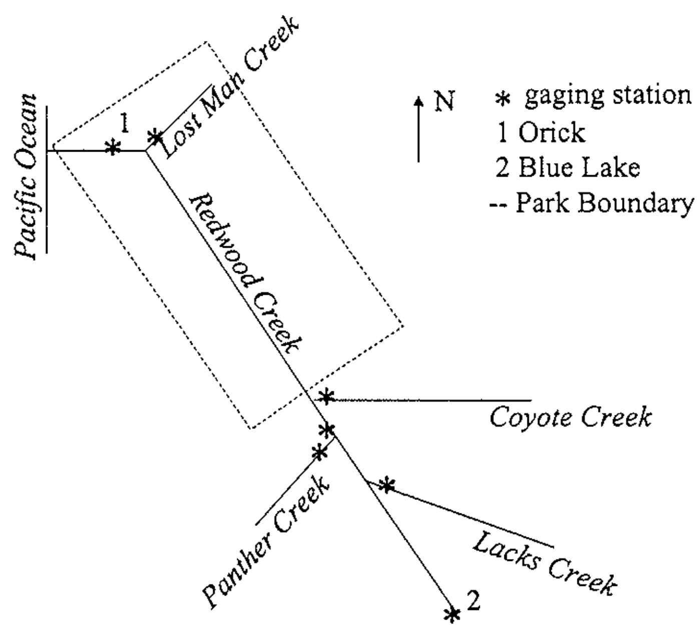

Figure 1. A Simple Representation of Redwood Creek

Table 1. List Of Catchment Regions and Respective Area[4]

Variable	Catchment Region	Area (sq. mi.)
s1	Coyote Creek	7.78
s2	Upper Redwood Creek	65.4
s3	Lacks Creek	16.9
s4	Panther Creek	6.07
s5	Lost Man Creek	3.46
s6	Lower Redwood Creek	115.8

4. The Upper Redwood Creek region consists of Redwood Creek above Panther Creek and below Blue Lake accounting for Lacks Creek. The Lower Redwood Creek region is downstream of Panther Creek accounting for Coyote Creek and Lost Man Creek.

METHODOLOGY

In this section we describe the methods used to in the research. The analysis starts with the construction of a theoretical NPS pollution control model, depicting the behavior of a budget-constrained manager who minimizes the pollution-related damage. Next, the SEF, which accommodates ill-posed data, is developed and applied to stream flow and ambient sediment loading data for Redwood Creek. Finally, the theoretical and methodological results are incorporated into a sediment control model for Redwood National Park. Here we examine the efficiency of a uniform pollution control policy versus a differentiated policies designed to capture the differences among the polluting sources. Sensitivity analysis is conducted with respect to the level and type of data collection to determine changes in the optimal treatment and data collection strategies.

A NONPOINT SOURCE POLLUTION MODEL

In the sediment control model, we assume the manager knows the location and size of the sources but there is uncertainty (incomplete information) about the generation of pollution from each source. The pollution control manager decides how to optimally allocate a limited budget between treatment effort and information acquisition. Treatment directly reduces the pollution generation. However, the manager can improve the treatment effectiveness by acquiring information that reduces the pollution generation uncertainty. If information is sequentially acquired (i.e., learning), then the degree of uncertainty declines with each sequential update. In some situations, a NPS problem may evolve to a PS one given that sufficient information is acquired.

We extend the control model in Farzin and Kaplan (2000) to a dynamic problem and incorporate information entropy as a measure of the pollution generation uncertainty. This new model generalizes the problem to capture any pollution type. In the model, information about the pollution generation is acquired through a single channel, namely data collection. In the context of sediment loading, data collection is often limited to stream flow for some sources and an aggregate sediment load measure for the waterway. If there is significant variability in and between the stream flow from the sediment loading sources, then changes in stream flow can be compared with changes in aggregate sediment load to estimate the unobservable pollution generated by each source (Kaplan 1999).

To construct the model let the states be p_i^n, the probabilities on the sediment loading states, and $s_i^n(t)$, the sediment loading states, where *i* denotes the state and *n* denotes the source. We set $i = 2$ to simplify the presentation and write the probabilities as p_1^n and $p_2^n = 1 - p_1^n$, which allows us to reduce the number of state equations from $n*i$ to n. The controls are $\alpha(t)$ and $R^n(t)$, the data collection intensity and level of treatment effort at each of the *n* source, respectively. The sediment loading uncertainty about the known sources at a given time is measured with the normalized information entropy metric over all sediment loading distributions or formally

$$I(p) = -(1/MaxEnt)\sum_n \sum_i p_i^n \ln p_i^n \ , \qquad (1)$$

where *MaxEnt* is the maximum entropy associated with the uniform distribution over the *p*'s. The state equations for the probabilities are

$$\dot{p}_1^n(t) = g^n(p_1^n(t), \alpha(t)), \forall n \qquad (2)$$

The sediment loading state equations are

$$\dot{s}_i^n(t) = -b^n R^n s_i^n, \forall i, \forall n \tag{3}$$

To reduce the number of state equations, we sum the n sediment loading state equations over i and obtain

$$\dot{s}_1^n(t) + \dot{s}_2^n(t) = -b^n R^n [s_1^n(t) + s_2^n(t)], \forall n \tag{4}$$

With the post-data probabilities and post-treatment sediment loading states we can construct the expected sediment loading from each source

$$E(s^n(t)) = \sum_i s_i^n(t) p_i^n(t) \tag{5}$$

We can interpret g^n as the change in $E(s^n)$ given the information acquired through data collection and $-b^n$ as the marginal decline in $E(s^n)$ given treatment effort.

In every time period the public manager is constrained by a budget $B(t)$, which is allocated between $\alpha(t)$ and $R(t) = (R^1(t), R^2(t), \cdots, R^N(t))$. We simplify the cost functions to focus our attention on the tradeoff between information acquisition and treatment effort. We assume costs are linear in data collection and treatment levels and treatment costs are identical across sites. Thus, the budget constraint is

$$B(t) = c\sum_n R^n(t) + \alpha(t)m \tag{6}$$

where c is the per unit treatment cost and m is the per unit data collection cost.

The optimal control problem for minimizing the damages from expected sediment loading at known sites that yields the optimal intensity of data collection and the optimal levels of treatment effort is

$$\underset{\alpha(t),R(t)}{Min} \quad \int_0^{\infty} e^{-rt} ED[S(t), P(t), \alpha(t), R(t)]dt \tag{7}$$

subject to

$$\dot{p}_1^n(t) = g^n(p_1^n(t), \alpha(t)), \forall n, \; p_1^n(t) = p_{1,0}^n, \forall n, \; p_2^n(t) = 1 - p_{1,0}^n, \forall n$$

$$\lim_{t\to\infty} p_1^n(t) = \overline{p}_1^n, \forall n, \; \lim_{t\to\infty} p_2^n(t) = 1 - \overline{p}_1^n, \forall n, \tag{7a}$$

$$\dot{s}_1^n(t) + \dot{s}_2^n(t) = -b^n R^n(s_1^n + s_2^n), \forall n, \; s_i^n(t) = s_{i,0}^n, \forall i, \forall n, \; \lim_{t\to\infty} s_i^n(t) = \overline{s}_i^n, \forall i, \forall n, \tag{7b}$$

$$E(s^n(t)) = \sum_i s_i^n(t) p_i^n(t), \tag{7c}$$

$$B(t) = c\sum_n R_n(t) + m\alpha(t). \tag{7d}$$

The expected damage function $ED(\cdot)$ is assumed to be twice continuously differentiable and convex in sediment loading. To solve this optimal control problem, we first construct the Hamiltonian,

$$\begin{aligned} H &= e^{-rt} ED(S(t), P(t), \alpha(t), R(t)) \\ &\quad + \sum_n \lambda_p^n g^n(p_1^n(t), \alpha(t)) - \sum_n \lambda_s^n b^n R^n(t)[s_1^n(t) + s_2^n(t)] \end{aligned}, \tag{8}$$

where λ_p^n and λ_s^n are the costate variables for the probabilities and sediment loading states at the n sources, respectively. Now we form the Lagrangian to incorporate the budget constraint, such that

$$\begin{aligned} L &= e^{-rt} ED(S(t), P(t), \alpha(t), R(t)) + \sum_n \lambda_p^n g^n(p_1^n(t), \alpha(t)) \\ &\quad - \sum_n \lambda_s^n b^n R^n(t)[s_1^n(t) + s_2^n(t)] + \mu\left[c\sum_n R^n(t) + \alpha(t)m - B(t)\right] \end{aligned}. \tag{9}$$

The necessary conditions to maximize the Lagrangian are

$$\frac{\partial L}{\partial \alpha} = e^{-rt} \frac{\partial ED}{\partial \alpha} + \sum_n \lambda_p^n \frac{\partial g^n}{\partial \alpha} + \mu m = 0, \tag{10a}$$

$$\frac{\partial L}{\partial R^n} = e^{-rt} \frac{\partial ED}{\partial R^n} - \lambda_s^n b^n \left[s_1^n(t) + s_2^n(t)\right] + \mu c = 0, \forall n, \tag{10b}$$

$$\dot{\lambda}_p^n = -e^{-rt} \frac{\partial ED}{\partial p_1^n} - \lambda_p^n \frac{\partial g^n}{\partial p_1^n}, \forall n, \ \lim_{t \to \infty} \lambda_p^n(t) = \overline{\lambda}_p^n, \forall n, \tag{10c}$$

$$\dot{\lambda}_s^n = -e^{-rt} \frac{\partial ED}{\partial [s_1^n(t) + s_2^n(t)]} + \lambda_s^n b^n R^n(t), \forall n, \ \lim_{t \to \infty} \lambda_s^n(t) = \overline{\lambda}_s^n, \forall n, \tag{10d}$$

$$\dot{p}_1^n(t) = g^n(p_1^n(t), \alpha(t)), \forall n, \ p_1^n(t) = p_{1,0}^n, \forall n, \ p_2^n(t) = 1 - p_{1,0}^n, \forall n, \tag{10e}$$

$$\lim_{t \to \infty} p_1^n(t) = \overline{p}_1^n, \forall n, \ \lim_{t \to \infty} p_2^n(t) = 1 - \overline{p}_1^n, \forall n, \tag{10f}$$

$$\dot{s}_1^n(t) + \dot{s}_2^n(t) = -b^n R^n [s_1^n(t) + s_2^n(t)], \forall n, \ s_i^n(t) = s_{i,0}^n, \forall i, \forall n,$$

$$\lim_{t \to \infty} s_i^n(t) = \overline{s}_i^n, \forall i, \forall n, \tag{10g}$$

$$B(t) = c \sum_n R_n(t) + \alpha(t) m, \ \mu \geq 0. \tag{10h}$$

The solution to these first order conditions defines the optimal levels of treatment and data collection through time.

THE ENTROPY FORMALISM

The use of information entropy in economic theory and econometric application is a recent phenomenon. Shannon (1948) first introduced the concept of information entropy as a measure of uncertainty. Khinchin (1957) formalized the mathematical foundations for the information-theoretic approach to uncertainty. The basic idea underlying the entropy approach is that the entropy of the uniform distribution (i.e., where any realization is as likely as another) is the maximum or upper bound on the information entropy metric. This represents complete uncertainty for a given random parameter and the case where there is no information except the upper and lower bounds of the uniform distribution. Conversely, a spiked distribution (i.e., only one realization can occur) has an entropy of zero. This lower bound corresponds to the case of certainty or complete information. Thus, normalizing the entropy measure by the maximum entropy bounds the uncertainty between zero and one and provides a bounded ordinal ranking of uncertainty. We can, in principle, model the entire spectrum of possible pollution control problems by incorporating this interpretation of information entropy into the management model.

There is a second application of the information entropy as well. That is, the estimation of parameter distributions. The entropy criterion selects the least informative distribution that is consistent with the data. Jaynes (1957) shows that this criterion maximizes the multiplicity. In this context, the selected distribution is the one that can be realized in the greatest number of ways, and consistent with what is known. The particular problem we face is estimating sediment-loading parameters for more sources than there are total ambient sediment measures. This is similar to the problem of reconstructing images from sparse data where maximum entropy methods have been used (Skilling 1989; Gull and Daniell 1978; and Gull and Skilling 1984).

There are two main advantages in using the entropy approach over least squares or maximum likelihood estimation that make it ideal for estimating NPS pollution parameters, where the underlying distributions for the natural system is unknown and the pollution generation is from diffuse sources. First, the error distribution does not have to be specified in the entropy approach. It is, however, estimated from the data. Second, given that there are more pollution source parameters to estimate than observations, the problem of inverting the matrix of explanatory variables is avoided with the entropy approach. This frees the researcher to make inferences about the model parameters without the degree of freedom restrictions that is necessary to employ least squares or maximum likelihood estimation.

In addition, to empirically incorporate learning (Bayesian updating), the SEF can be employed to update parameter distributions as new data become available (Kaplan and Howitt 2000). Formal modeling of Bayesian learning by integrating over joint density functions is often very cumbersome (Fahrmeir 1992). Fortunately, The SEF is equivalent to Bayes' theorem (Kaplan and Howitt 2000; Zellner 1988; Golan, Judge and Miller 1996). Given that the SEF is easier to empirically model and has desirable consistency properties, we employ it in estimating the unknown sediment loading parameters.

Previous empirical application of the entropy metric in state space modeling, which is used to derive the sequential entropy filter approach, has been limited. Golan, Judge and Karp (1996), in the context of a dynamic discrete time model, present the generalized maximum entropy (GME) approach for reconstructing unobservable state space parameters. Similarly, Fernandez (1997) reconstructs the state and observation parameters for a dynamic wastewater treatment plant model using a GME approach. Neither of these studies considers the sequential updating potential of the cross entropy approach in dynamic models.

THE SEQUENTIAL ENTROPY FILTER

In the empirical estimation a single ambient sediment measure is disaggregated among the six catchment regions and then sequentially updated, as new data becomes available (i.e., with each observation). The random state equation parameters are also estimated and sequentially updated. Two estimation scenarios are considered. In the first case, defined as high intensity data collection, we use the disaggregated stream flow data. In the second case, defined as low intensity data collection, the stream flow data for the lower reach of the creek (i.e., downstream from the Redwood Creek at Panther Creek gaging station) are combined into one stream flow measure, and the stream flow data for the upper reach of the creek are combined into a separate single measure. We then use the relative area of each catchment region to reapportion the stream flow data to each region. These different data collection scenarios provide estimates of the sediment loading distributions, which are used later in the simulated sediment control model.

Let the empirical state and observation equations be defined as

$$E(s^n(t+1)) = \left(g^n(\alpha(t+1)) - \beta^n(\alpha(t+1))R^n(t)\right)E(s^n(t)) + v^n(t), \forall n \qquad (11)$$

$$\ln q^n(t) = E(s^n(t))\ln(flw^n(t)) + w^n(t), \forall n \qquad (12)$$

where v^n and w^n are random errors which are also estimated. Equation (11) is a simplified combination of equations (2) and (3) with an additive error term. We derive equation (12) by adding an error term to the log transformation on a known hydrological relationship between stream flow and sediment loading (Singh and Krstanovic 1987), which is

$$Q(t) \equiv \sum_n q^n(t) \equiv \sum_n \overline{q}^n(t) = \sum_n (flw^n(t))^{E(s^n(t))} . \qquad (13)$$

Q is the ambient sediment load measure, $q^n(t)$ is the actual unobservable sediment generated from the *nth* source, $\overline{q}^n(t)$ is the expected sediment generated from the *nth* source, and $flw^n(t)$ is the known stream flow from the *nth* source.[5]

To estimate probability distributions for the random model parameters, the following reparameterization from parameter space to probability space is necessary.

$$E(s^n(t)) = \sum_i s_i^n(t) p_i^n(t), \forall n, \; g^n(t) = \sum_i z_i^{n,g} \gamma_i^n(t), \forall n, \; b^n(t) = \sum_i z_i^{n,b} \beta_i^n(t), \forall n,$$

$$w^n(t) = \sum_i z_i^{n,w} p_i^{n,w}(t), \forall n, \text{ and } v^n(t) = \sum_i z_i^{n,v} p_i^{n,v}(t), \forall n, \tag{14}$$

where $s_i^n(t), z_i^{n,g}, z_i^{n,b}, z_i^{n,w}$, and $z_i^{n,v}$, are support values for the respective distributions that represent the constraints on the probability space. The probability distributions sum to unity;

$$\sum_i p_i^n(t) = 1, \forall n, \; \sum_i \gamma_i^n(t) = 1, \forall n, \; \sum_i \beta_i^n(t) = 1, \forall n,$$

$$\sum_i p_i^{n,w}(t) = 1, \forall n \text{ and } \sum_i p_i^{n,v}(t) = 1, \forall n \tag{15}$$

In each time period t, the entropy specification of the objective function is

$$\begin{aligned} Min \sum_n \sum_i p_i^n \ln\left(p_i^n \big/ \tilde{p}_i^n\right) + \sum_n \sum_i \beta_i^n \ln\left(\beta_i^n \big/ \tilde{\beta}_i^n\right) \\ + \sum_n \sum_i \gamma_i^n \ln\left(\gamma_i^n \big/ \tilde{\gamma}_i^n\right) + \sum_n \sum_i p_i^{n,w} \ln p_i^{n,w} + \sum_j p_i^{n,v} \ln p_i^{n,v} \end{aligned} \tag{16}$$

subject to equations (11) - (15), where $\tilde{p}_i^n \equiv p_i^n(t-1), \forall n$, $\tilde{\beta}_i^n \equiv \beta_i^n(t-1), \forall n$, and $\tilde{\gamma}_i^n \equiv \gamma_i^n(t-1), \forall n$ are the prior probabilities.[6]

---

5. The identity on the left-hand side of equation (14) formalizes the case that the observable ambient load is equivalent to the sum of sediment from each source and to the sum of expected sediment from each

We assume the initial prior probabilities $\tilde{p}$, $\tilde{\beta}$, and $\tilde{\gamma}$, are uniformly distributed. In other words, any state or value is as likely as any other and thus the probabilities will be equal to $\frac{1}{\sum i}$, where $\sum i$ is the total number of support values.[7] For tractability, and to confer with the earlier theoretical discussion, the number of support values is limited to two. The support values for s^n and $z^{n,b}$ range from 0.0 to 2.0. This implies that rainfall has a non-negative impact on sediment loading and treatment has a non-positive impact on sediment loading. Further, the upper bounds on s^n are set to contain estimates derived for other river systems as reported in Singh and Krstanovic (1987). The estimation will inform us as to which catchment region deviates from the original uniform prior distribution. The values for $z^{n,g}$ range from 0.0 to 2.0 as well. This range allows the information effect (g^n) to increase or decrease the expected sediment loading but does not allow the loading to decrease with increases in rainfall. The values for z^w, and z^v, range from –50.0 to 50.0, and from –10.0 to 10.0 respectively. These ranges are set well beyond the plausible values for the estimated parameter. After estimating posterior distributions from the first observation, all subsequent prior probabilities are the previous period estimated posterior probabilities.

SIMULATION ANALYSIS

The control model defined in equation (7) was calibrated with the estimation results reported above and additional data provided by the Redwood National Park staff. The cost

---

source.

6. We drop the argument t from the objective function for convenience.

function coefficient was estimated with least squares using a linear specification without an intercept term such that the value of the estimated cost coefficient is 2178.1 with a t-statistic of 13.683 and a R^2=0.34. The damage function was estimated based upon the assumption that the sediment related damage is greater than or equal to the resources spent on sediment control. The estimated damage function in dollars is

$$\ln(\text{Damage}) = \underset{(-2.839)}{-7.253} + \underset{(8.054)}{1.674}\ \ln(\text{sediment}),\ R^2{=}0.84$$

where sediment is measured in cumulative average daily tons per rain season and the t-statistic are reported below the estimated parameters. The data collection cost under the varying intensities are m(low intensity) = \$41,154 and m(high intensity) = \$93,712, and the annual budget is B = \$200,000.

The policy analysis involves simulating the model under three scenarios. First, a uniform treatment policy is simulated, where all catchment regions are assumed to load sediment in proportion to their relative area. This first scenario coincides with best management practices mentioned earlier in the paper. Second, two scenarios depicting low and high intensity data collection are simulated, where the manager incorporates the information acquired through data collection to reallocate resources from relative cleaner regions to relatively dirtier regions.

PRINCIPAL FINDINGS AND SIGNIFICANCE

THEORETICAL MODEL

The control problem does not allow us to solve for an analytical solution; however, we can still derive some qualitative properties from the necessary conditions. First, equations (10a) and

---

7. Although the managers have knowledge to differentiate the sources, the uniform distribution is chosen to reflect the knowledge prior to the enactment of the sediment control program.

(10b) provide the optimal conditions for choosing data collection intensities and treatment effort. These optimal conditions tell the familiar story of selecting inputs so that the expected present value of marginal benefits from employing an input equals it's marginal cost. Farzin and Kaplan (2000) noted that information acquisition is a collective good and thus equation (10a) equates the expected marginal reductions in sediment loading over all sources with the marginal opportunity cost of acquiring information. There is another interesting result obtained from (10a) and (10b) as well. We solve (10a) and (10b) for the shadow value on the budget constraint (μ) and then substituting for μ yields

$$
\begin{aligned}
1/m\left(e^{-rt}\frac{\partial ED}{\partial \alpha}+\sum_{n}\lambda_p^n\frac{\partial g^n}{\partial \alpha}\right) &= 1/c\left(e^{-rt}\frac{\partial ED}{\partial R^1}-\lambda_s^1 b^1\left[s_1^1(t)+s_2^1(t)\right]\right) \\
&= \cdots = 1/c\left(e^{-rt}\frac{\partial ED}{\partial R^N}-\lambda_s^N b^N\left[s_1^N(t)+s_2^N(t)\right]\right)
\end{aligned}
\quad (17)
$$

Equation (17) shows that the optimal allocation of data collection and treatment effort over time will equate the ratio of expected present value of marginal benefits to marginal cost over all inputs. Thus, the manager will equate the adjusted marginal benefits obtained from a collective good (data collection) with the adjusted marginal benefits from the treatment effort (site-specific good).

ESTIMATION RESULTS

The estimation results for the sequential entropy filter appear in Figures 2, 3 and 4. Figure 2 illustrates the information acquired under the two data collection intensities. It appears that more information is acquired (uncertainty is reduced) in the high-intensity policy than in the low-intensity policy. Over the time horizon, roughly twenty-five to thirty percent of the uncertainty associated with the pollution generation is reduced when data is collected and information is acquired.

Figures 3 and 4 illustrate the entropy estimates for the sediment loading parameter for each of the six catchment regions. The H and L listed in the parentheses represents the high and low-intensity data collection, respectively. The estimation results appear rather similar for high and low intensity data collection. We can infer from these results that the larger catchment regions (s2, s3 and s6) contribute less sediment per square mile than the smaller regions (s1, s4 and s5) given the relatively smaller estimated sediment loading parameter for the larger regions. We also see from Figure 5 and 6 that estimate for s6 is much more responsive to learning than the other estimates as indicated by the volatility of the g parameters in both the high and low-intensity data collection scenarios. This may be attributable to the fact that treatment only occurs in this region and the influence of treatment on region s6 may be captured by the g parameter. This might follow from the estimated b parameter, which remains at the initial prior distribution (with an expected value of 0.001) throughout the sequential estimation, implying that any changes in s6 due to treatment may be picked up in the g parameter.

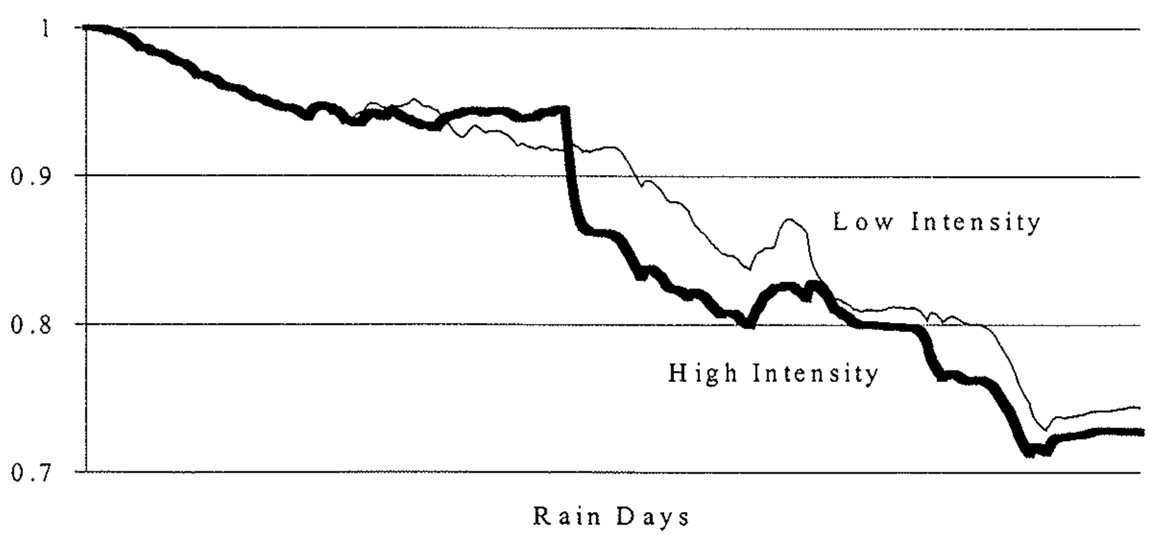

Figure 2. Normalized Entropy for Low and High Intensity Data Collection.

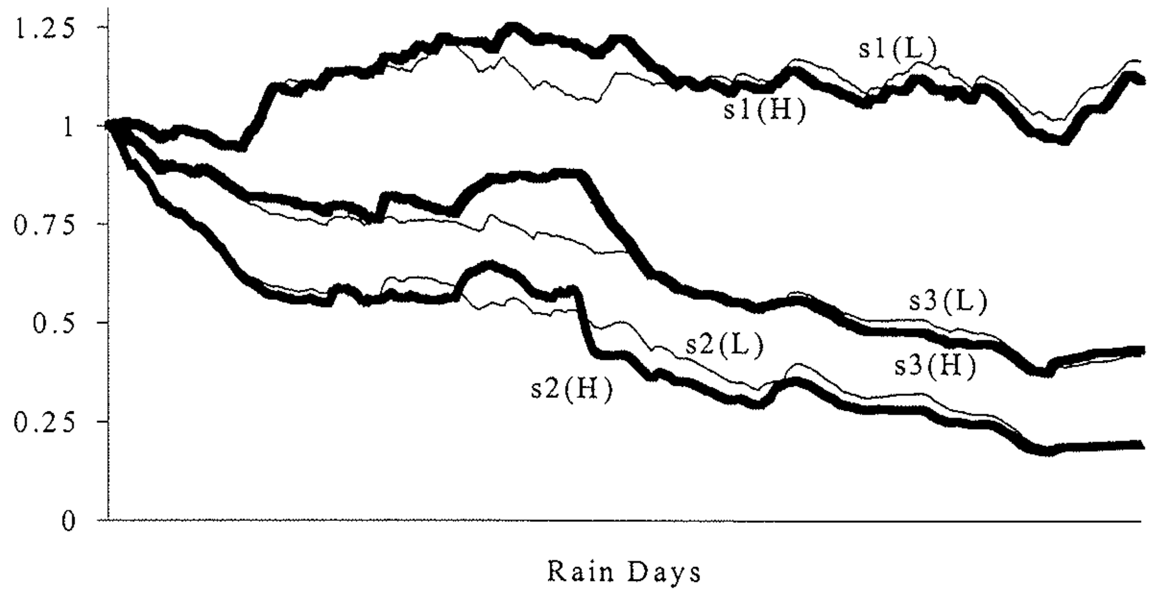

Figure 3. Estimated Sediment Loading Parameters s1, s2, and s3.

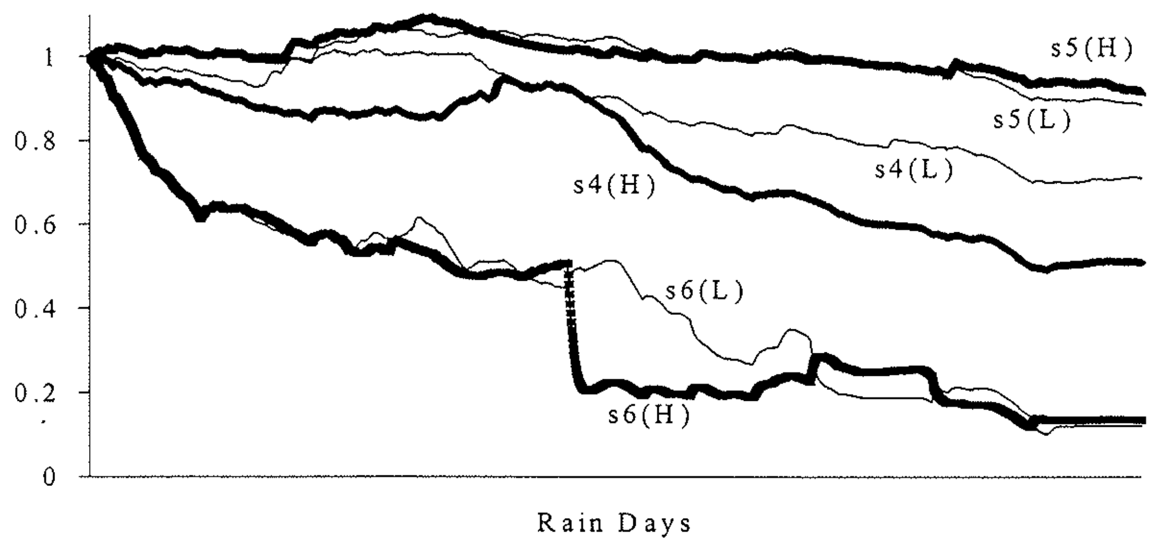

Figure 4. Estimated Sediment Loading Parameters s4,s5,and s6.

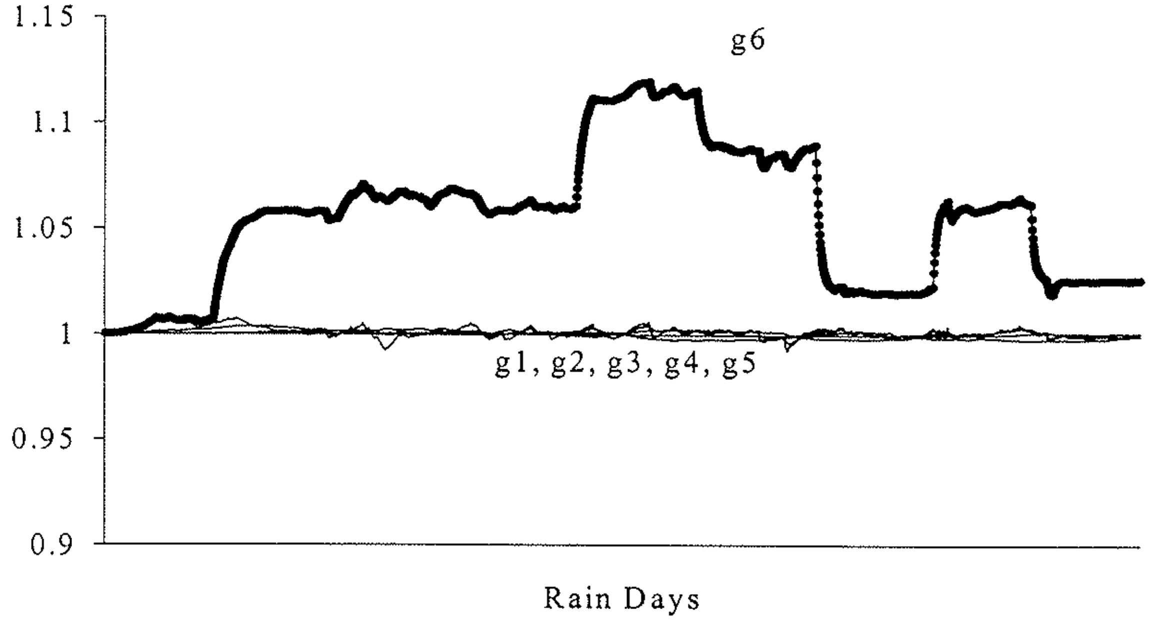

Figure 5. Estimated g Parameters under Low Intensity Data Collection.

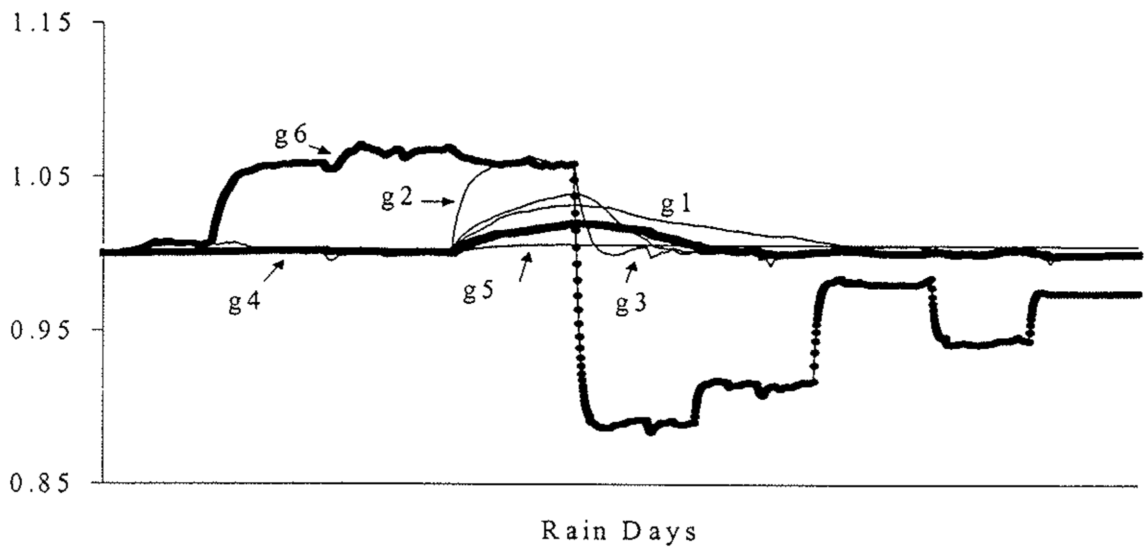

Figure 6. Estimated g Parameters under High Intensity Data Collection.

SIMULATION RESULTS

The results from the policy simulations are shown in Figure 7. We see that the manager improves the treatment effectiveness and lowers damages in all rain seasons when the high and low-intensity data collection policies are chosen. Interestingly, the low-intensity data collection policy is the optimal strategy for minimizing sediment related damages. This suggests that the cost of data collection adversely influences the high-intensity data collection strategy so that resources are limited and the cheaper, low-intensity data collection leaves more resources to treatment and thus lower damages. The optimal treatment effort under the high and low-intensity data collection is listed in Table 2. The optimal treatment effort under the BMP policy is on average 19.9 for Region s2 and 71.92 for Region s6. It is clear that the gain in information and thus treatment effectiveness under the high and low-intensity data collection policy translates into lower sediment related damage, although less budgetary resources are available for treatment given the data collection expense.

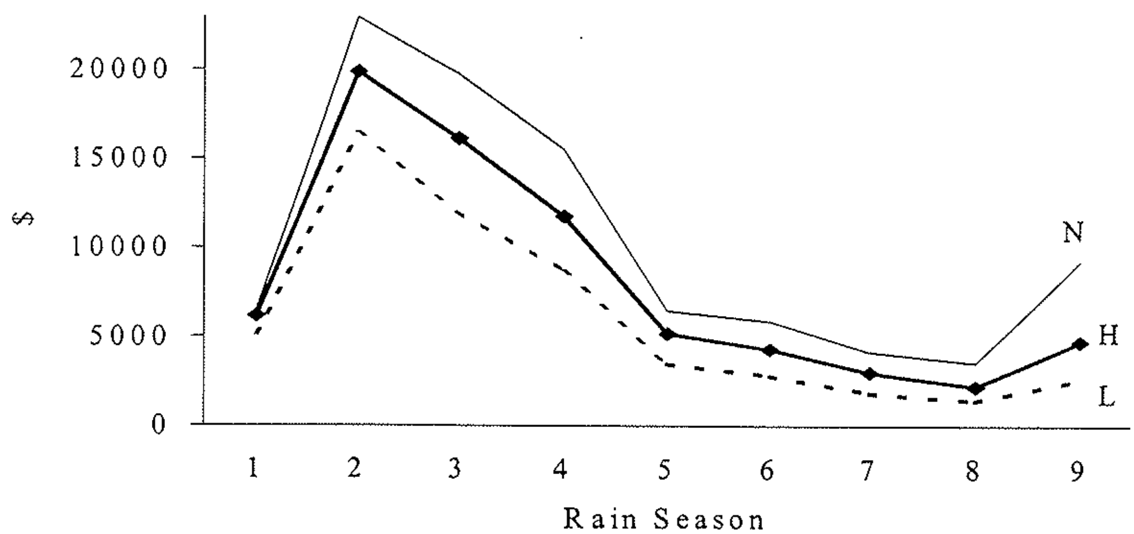

Figure 7. Estimated Damages under Three Policy Scenarios: (N) no data collection; (H) High Intensity Data Collection; and (L) Low Intensity Data Collection.

Table 2. Optimal Treatment under (H) High and (L) Low Intensity Data Collection.

Rain	High Intensity Data Collection						Low Intensity Data Collection					
Season	s1(H)	s2(H)	s3(H)	s4(H)	s5(H)	s6(H)	s1(L)	s2(L)	s3(L)	s4(L)	s5(L)	s6(L)
1	23.1	-	25.7	-	-	-	42.5	-	15.3	-	15.2	-
2	11.3	-	18.1	19.4	-	-	9.1	-	30.0	11.4	22.4	-
3	15.9	-	17.7	15.2	-	-	32.2	-	5.0	4.8	30.9	-
4	29.6	-	19.2	-	-	-	-	-	35.8	37.1	-	-
5	3.7	-	13.9	22.9	8.3	-	18.4	9.7	21.9	13.8	9.1	-
6	25.9	-	10.3	12.6	-	-	13.2	16.7	14.4	14.9	13.8	-
7	8.0	8.5	12.9	9.9	9.5	-	-	25.1	24.4	23.4	-	-
8	9.8	-	17.7	3.0	18.3	-	22.1	-	-	-	18.2	32.6
9	40.7	-	-	0.8	7.3	-	28.2	-	-	1.1	37.4	6.2

DISCUSSION

This research examines the role of information in the control of nonpoint source (NPS) pollution. The theoretical model and empirical method developed in the analysis brings to light the importance of information acquisition in controlling NPS pollution and in statistically estimating pollution loading parameters from NPS pollution data. In addition, the empirical policy analysis allows us to show the usefulness of information acquisition for an actual NPS problem in reducing uncertainty and in producing lower damages than would otherwise occur.

The theoretical analysis, which incorporated both information acquisition and a fiscal constraint, highlights the trade off between treatment effort and treatment effectiveness. We explicitly consider the opportunity cost of acquiring information, in the case of reduced treatment effort due to exogenous budget restrictions. We analyze the problem within a budget constrained management framework to focus attention on a more practical setting for studying NPS pollution control. The results show how the manager can exploit the heterogeneity of the sites to reduce expected damages with lower treatment expenditures. With sufficient information, the NPS pollution manager can move away from uniform treatment policies, such as best management practices, and focus on policies that exploit the heterogeneity of the landscape and of the pollution generation across sources. Furthermore, if it is economically feasible for the manager to acquire the information on NPS generation (i.e., reduce the pollution generation uncertainty), the problem evolves towards one with point source characteristics.

Information acquisition also defines the methodological component of the research. The use of the entropy formalism, which is the basis for information-theoretic econometrics, allows us to draw inferences from undersized samples where other techniques cannot yield estimates without restrictive assumptions. The restrictive assumptions necessary to employ least squares or maximum likelihood approaches induce potential biases in these approaches when confronted with small or undersized samples. We present the entropy approach for optimally updating parameter estimates when the data comes from an undersized sample. Furthermore, we argue that the entropy model has several advantages for estimating parameters from undersized samples or when the error distributions are unknown. For instance, when the data is not gaussian, the entropy specification provides a unique solution, whereas the TGME and KF are no longer applicable. In addition, the entropy specification can be used to solve well or ill-posed data processes. When the data are ill-posed, this specification for estimating optimal information acquisition imposes less restrictive, and more realistic, assumptions than that of the TGME approach. Finally, the entropy model is computationally simple and provides a method for readily calculating empirical Bayesian estimates.

In the empirical application, the sequential entropy filter is adapted to the case of sediment loading in Redwood Creek, which flows into and through Redwood National Park. We employ the theoretical and methodological methods presented above to conduct empirical Bayesian analysis. The results indicated that greater intensity of data collection produces greater reductions in uncertainty. In other words, the problem evolves more quickly towards a point source problem as the intensity of data collection increases. This is a result that depends on the particular case being analyzed. It should be made clear that some NPS problems may never evolve to PS ones.

In addition, we present a simulated policy analysis, where three policy options were considered. First, a uniform treatment policy is considered. Under this policy no data is collected and all catchment regions are assumed to produce the same sediment per square mile. This policy is analogous to the best management practices policy currently promoted. The other policies considered in the analysis included a high and a low-intensity data collection policy, where the manager collects data at varying intensities in order to improve the treatment effectiveness and thus reduce sediment loading within the creek. The simulation results show that when low-intensity information is acquired, the manager can more efficiently allocate limited resources and lower the sediment loading, which threatens salmon habitat and the Tall Trees Grove. This empirical analysis also indicated that when the high and low-intensity information was acquired, and uncertainty reduced, it was advantageous to move away from a uniform best management strategy to a heterogeneous policy, which exploits the differences among the polluting sites.

REFERENCES

Farzin, Y.H. and J.D. Kaplan. "Nonpoint Source Pollution Control under Incomplete and Costly Information," *Fondazione Eni Enrico Mattei (FEEM)* Working Papers, No.32.99, May 1999.

Farzin, Y. H., "Optimal Pricing of Environmental and Natural Resource Use with Stock Externalities," *Journal of Public Economics* **62**(1996):31-57.

Fernandez, L. "Estimation of Wastewater Treatment Objectives through Maximum Entropy." *Journal of Environmental Economics and Management* **32**:2(1997):293-308.

Golan, A, Judge, G and D. Miller. *Maximum Entropy Econometrics: Robust Estimation with Limited Data*. John Wiley and Sons: New York, 1996.

Golan, A. Judge G., and L. Karp. "A Maximum Entropy Approach to Estimation and Inference in Dynamic Models or Counting Fish in the Sea Using Maximum Entropy." *Journal of Economic Dynamics and Control* **20**(1996):559-582.

Gray, W.B., and M.E. Deily. "Compliance and Enforcement: Air pollution Regulation in the U.S. Steel Industry." *Journal of Environmental Economics and Management* **31**(1996):96-111.

Gull, S.F. and J. Skilling. "Maximum Entropy Method in Image Processing." *IEEE Proceedings* **131**:6(October 1984) 646-659.

Gull, S.F., and G.J. Daniell. "Image Reconstruction from Incomplete and Noisy Data." *Nature* **272**:20(April 1978) 686-690.

Jaynes, E.T. "Information Theory and Statistical Mechanics." *Physics Review* **106**(1957):620-30.

Kaplan, J.D. "Nonpoint Source Pollution Control, Incomplete Information and Learning: An Entropy Approach," Ph.D. Thesis, Department of Agricultural and Resource Economics, University of California, Davis (1999).

Kaplan, J.D. and R.E. Howitt, "An Entropy Filter Approach to Sequential Bayesian Estimation." Working Paper, Department of Agricultural and Resource Economics, University of California, Davis January 2000.

Khinchin, A.I. *Mathematical Foundations of Information Theory*. Dover Publications: New York,1957.

Mount, J.F. *California Rivers and Streams: The Conflict Between Fluvial Process and Land Use*. University of California Press: Berkeley, 1995.

Segerson, K. "Uncertainty and Incentives for Nonpoint Source Pollution Control." *Journal of Environmental Economics and Management* **15**(1988):87-98.

Shannon, C.E. "A Mathematical Theory of Communications." *Bell System Technical Journal* **27**(1948):379-423.

Singh, V. P. and P.F. Krstanovic. "A Stochastic Model for Sediment Yield Using the Principle of Maximum Entropy." *Water Resource Research* **23**:5(1987):781-793.

Skilling, J. Ed., *Maximum Entropy and Bayesian Methods*. Kluwer: Cambridge, 1989.

Spreiter, T.A., Franke, J.F. and D.L. Steensen. "Disturbed Lands Restoration: The Redwood experience." Paper Presentation at the Eight Biennial Conference on Research and Resource Management in Parks and Public Lands, George Wright Society, Hancock, Michigan, April 17-21, 1995.

United States Department of the Interior (DOI). *Watershed Rehabilitation Plan: Redwood National Park*, United States National Park Service, US Government Press, Washington DC, 1981.

United States Environmental Protection Agency (EPA). Chapter 3: Management Measures for Forestry-Introduction, http://www.epa.gov/OWOW/NPS/MMGI/ Chapter3/ch3-1.html, 1999.

United States General Accounting Office (GAO). *Water Quality: Federal Role in Addressing-and-Contributing to-Nonpoint Source Pollution*, GAO/RCED-99-45, US Government Press, Washington DC, 1999.

Zellner, A. "Optimal information Processing and Bayes' Theorem." *American Statistician* **42**(1988):279-84.